Prosecutorial Use of Generative Artificial Intelligence in Criminal Proceedings: Professional Responsibility, Evidentiary Integrity, and the Limits of Delegated Legal Judgment

Sergio Pommier Gallo, PhD

ORCID

0009-0000-7387-3943

Web of Science ResearcherID

QDO-1840-2026

Abstract

The increasing use of generative artificial intelligence (GAI) by prosecutors, investigators, and criminal-justice practitioners introduces a distinct legal problem that remains insufficiently differentiated from broader debates concerning algorithmic decision-making: the professional and procedural consequences of delegating elements of prosecutorial analysis, drafting, evidence review, and legal reasoning to generative systems. Existing scholarship has focused substantially on algorithmic prediction, automated policing, judicial decision-making, and AI-generated evidence, while comparatively less attention has been devoted to the professional responsibility consequences arising when prosecutors themselves employ generative systems in the preparation and presentation of criminal cases. This article examines the legal limits of prosecutorial reliance on GAI through doctrinal and comparative analysis of professional responsibility, disclosure, evidentiary integrity, confidentiality, candor toward tribunals, and prosecutorial duties of fairness. Particular attention is given to the American Bar Association's Formal Opinion 512 on generative artificial intelligence,

Model Rule 3.8, the United States Department of Justice's 2024 Artificial Intelligence and Criminal Justice Final Report, the European Union Artificial Intelligence Act, and Council of Europe standards concerning human rights, accountability, transparency, and human oversight. The analysis demonstrates that the central legal issue is not whether prosecutors may use AI-assisted tools, but whether such use can remain subordinate to professional judgment, evidentiary verification, and institutional responsibility. The article argues that GAI may legitimately assist prosecutorial work when used as an instrument rather than a delegated decision-maker; however, its outputs cannot independently establish factual propositions, legal authorities, evidentiary conclusions, or charging judgments. A defensible legal framework therefore requires human verification, traceable professional responsibility, preservation of relevant records where legally necessary, protection of confidential information, and disclosure of AI-related information when its omission would impair the defendant's ability to test the reliability or significance of material evidence.

Keywords

generative artificial intelligence; criminal prosecution; prosecutorial ethics; professional responsibility; criminal evidence; disclosure; evidentiary integrity; legal technology; human oversight; criminal procedure

Introduction

Generative artificial intelligence has moved rapidly from experimental technology into professional legal environments. Unlike earlier forms of legal automation, generative systems do not merely retrieve information or classify documents. They generate text, summarize records, formulate arguments, identify apparent authorities, organize factual material, and produce outputs that can resemble conventional legal analysis. Their increasing accessibility creates significant opportunities for criminal-justice institutions, particularly in environments characterized by substantial documentary and evidentiary burdens.

The use of GAI by prosecutors, however, raises a problem that cannot be resolved merely through general principles concerning judicial use of artificial intelligence. The

prosecutor occupies a distinctive institutional position. The prosecutor is not simply an advocate seeking to maximize the probability of success. Professional standards impose obligations concerning probable cause, disclosure, fairness, candor, and the treatment of exculpatory information. The use of a generative system therefore occurs within an institutional framework in which mistakes can affect liberty, charging decisions, disclosure obligations, plea negotiations, and the presentation of evidence.

The American Bar Association's Formal Opinion 512 recognizes that lawyers using generative AI remain subject to professional duties concerning competence, confidentiality, communication, supervision, meritorious claims, candor toward the tribunal, and fees. It does not treat technological adoption as a substitute for professional judgment.

The problem is consequently more specific than the proposition that "AI can make mistakes." Legal systems already possess doctrines governing unreliable evidence, professional negligence, disclosure violations, misleading submissions, and improper delegation. The question is how these existing principles should apply when a prosecutor incorporates a probabilistic generative system into the preparation of a criminal case.

This article does not propose an autonomous theory of AI prosecution. Instead, it examines whether existing principles of professional responsibility and criminal procedure are sufficiently capable of governing prosecutorial use of GAI and identifies where additional doctrinal clarification is necessary. The analysis proceeds from a conservative premise: generative AI should remain an instrument of professional activity, while the legally consequential judgments produced from its use must remain attributable to a competent human prosecutor.

The issue is particularly important because the distinction between administrative assistance and substantive legal judgment can become difficult to maintain. A system used to summarize discovery may influence which evidence a prosecutor notices. A system used to draft an indictment may influence the factual characterization of conduct. A system used to identify allegedly exculpatory material may influence disclosure decisions. A system used to research case law may generate authorities that

appear authentic but require verification. The technological operation may therefore be formally assistive while substantively influencing legal judgment.

The resulting problem is one of professional responsibility, evidentiary integrity, and institutional accountability.

Research Question

To what extent may prosecutors employ generative artificial intelligence in criminal proceedings without violating existing duties of professional competence, evidentiary integrity, disclosure, confidentiality, candor, and individualized human judgment?

Hypothesis

Existing doctrines of professional responsibility and criminal procedure can accommodate prosecutorial use of generative artificial intelligence provided that GAI remains subordinate to independent human judgment and does not become an unverified source of factual, legal, or evidentiary conclusions. Where AI-generated outputs materially influence prosecutorial decisions or the presentation of evidence, the prosecutor's obligations of verification, supervision, confidentiality, and disclosure may require heightened procedural safeguards.

Theoretical Framework

1. Prosecutorial Authority as a Professionally Constrained Function

The prosecutor occupies a position that differs fundamentally from that of an ordinary litigant. The prosecutor possesses institutional authority to initiate or continue criminal proceedings, negotiate dispositions, determine evidentiary priorities, and make representations to the court. The exercise of that authority is therefore constrained by professional and procedural duties.

ABA Model Rule 3.8 identifies special responsibilities of prosecutors, including the obligation not to prosecute charges unsupported by probable cause and the obligation to

make timely disclosure to the defense of evidence or information known to the prosecutor that tends to negate guilt or mitigate the offense.

These obligations become technologically significant when prosecutors use GAI to process information. If a generative system summarizes a large evidentiary record and omits an exculpatory item, the prosecutor cannot transfer responsibility for that omission to the system. The professional duty remains attached to the human institution exercising prosecutorial authority.

The same principle applies to legal research. If GAI generates a non-existent case, misstates a precedent, or attributes an incorrect proposition to an authentic authority, the prosecutor remains responsible for the resulting representation. Generative output is therefore epistemically subordinate to professional verification.

2. GAI as an Assistive Rather Than Substitutive Technology

The distinction between assistance and substitution provides the central conceptual boundary.

An assistive technology performs a task that remains subject to meaningful human evaluation. A substitutive technology effectively determines the result without adequate independent verification.

The European Union Artificial Intelligence Act provides an important comparative reference. Article 14 requires effective human oversight for high-risk AI systems and links such oversight to the prevention or minimization of risks to fundamental rights. The Regulation further emphasizes that persons responsible for oversight must possess appropriate competence, training, and authority.

Although the precise application of the AI Act depends on the function and legal context of a particular system, its regulatory architecture illustrates a broader principle: human oversight is not satisfied merely by placing a human somewhere in the workflow. The human must possess the capacity and authority to understand, evaluate, reject, or otherwise control the system’s output.

Applied to prosecution, this means that a prosecutor cannot legitimately treat an AI-generated conclusion as presumptively correct merely because the output was produced by an advanced system. The prosecutor must retain independent judgment.

3. Epistemic Reliability and Legal Responsibility

Generative AI presents a distinctive epistemic problem. Traditional legal research tools generally return existing documents. A generative model can produce an apparently coherent proposition without establishing that the proposition corresponds to an actual source.

This distinction is particularly important in criminal proceedings because legal validity depends upon traceability. A factual assertion must be connected to evidence. A legal proposition must be connected to an authoritative source. An expert conclusion must be connected to an identifiable methodology and qualified person. A prosecutorial representation to a tribunal must be capable of professional verification.

The problem is therefore not simply technological inaccuracy. It is the possibility of **epistemic displacement**, whereby the appearance of reasoned output obscures the absence of a legally sufficient foundation.

The ABA's Formal Opinion 512 emphasizes that lawyers must understand the capabilities and limitations of GAI and must satisfy existing professional obligations when using it.

Accordingly, the relevant legal standard should not be whether a GAI output is generally accurate. The relevant standard is whether the prosecutor has independently established the accuracy and legal sufficiency of the particular proposition being relied upon.

Methodology

This article employs qualitative doctrinal legal research, comparative legal analysis, and normative analysis.

The doctrinal component examines professional responsibility principles, criminal procedural obligations, disclosure requirements, human-oversight standards, and emerging AI regulation.

The comparative component considers developments in the United States, the European Union, and the Council of Europe. The United States provides a useful framework through professional responsibility rules and federal criminal discovery principles. The European Union provides an emerging regulatory model through the AI Act. The Council of Europe provides a broader human-rights and rule-of-law framework applicable to AI throughout its lifecycle.

The normative component evaluates whether existing legal principles adequately allocate responsibility when generative AI becomes integrated into prosecutorial workflows.

The analysis does not rely on empirical claims regarding error rates or predictive performance that cannot be independently established within the doctrinal record. The focus is instead on legal responsibility, procedural safeguards, and institutional accountability.

Results and Legal Analysis

1. The Duty of Verification

The first and most fundamental requirement is independent verification.

A prosecutor may use GAI to generate a preliminary memorandum, summarize documents, identify potential legal issues, organize discovery, or suggest lines of inquiry. None of these uses should relieve the prosecutor of the duty to verify consequential outputs.

This principle follows from existing professional obligations rather than requiring a new doctrine. Professional competence requires the lawyer to understand the technology sufficiently to identify its limitations. Formal Opinion 512 expressly addresses this requirement in relation to GAI.

In criminal proceedings, the standard should be particularly demanding because errors can directly affect liberty.

Verification should therefore occur at three levels.

First, **source verification** requires confirmation that cited cases, statutes, regulations, articles, and other authorities actually exist and support the proposition attributed to them.

Second, **factual verification** requires confirmation that factual assertions correspond to the underlying evidentiary record.

Third, **procedural verification** requires confirmation that the proposed legal action is compatible with applicable procedural requirements.

A generated memorandum may therefore function as a starting point, but it cannot become the evidentiary or legal foundation of a prosecutorial decision without human validation.

2. AI-Assisted Discovery and Exculpatory Material

The most consequential application may be AI-assisted review of discovery.

Modern criminal cases can contain enormous quantities of emails, messages, photographs, video, financial records, digital files, and other information. AI-assisted review can help identify potentially relevant material. However, a system's failure to identify exculpatory information cannot eliminate the prosecutor's disclosure obligations.

Rule 3.8 recognizes a prosecutor's responsibility concerning evidence that tends to negate guilt or mitigate an offense.

The United States Federal Rules of Criminal Procedure also illustrate the importance of access to material relevant to the defense. Rule 16 provides for disclosure and inspection of specified documents, data, objects, scientific-test results, and expert-related information subject to the rule's conditions.

The legal problem is therefore not whether AI can assist with discovery. It plainly can. The problem is whether an institution can rely on an automated or generative process without maintaining sufficient human supervision to ensure that legally material information is not systematically excluded.

A defensible approach requires treating AI-assisted review as a method of organizing information, not as the final determinant of what is legally significant.

3. Generative AI and Legal Research

Legal research presents a related but distinct risk.

A prosecutor using GAI to identify authorities may receive:

- authentic authorities accurately summarized;
- authentic authorities inaccurately summarized;
- authorities that exist but do not support the proposition asserted;
- outdated authorities;
- jurisdictionally irrelevant authorities;
- or entirely fabricated authorities.

The last category is particularly problematic because the output may appear professionally plausible.

The legal response should not be to prohibit AI-assisted research. Such a prohibition would unnecessarily prevent useful technology from supporting professional work. The appropriate response is to preserve conventional source-verification requirements.

Every authority used in a charging document, memorandum, motion, brief, or oral submission should therefore be independently located and examined by the prosecutor or a competent member of the legal team.

Generative AI may identify a possible authority.

It should not authenticate its own authority.

4. Drafting Charging Documents

The use of GAI to draft indictments, informations, complaints, or other charging documents raises a more serious issue because drafting can influence substantive prosecutorial judgment.

A charging document is not merely an administrative text. It expresses the state's formal allegations against an individual and may determine the legal and factual boundaries of the prosecution.

A system that generates language can inadvertently transform tentative factual information into apparently established allegations. It can also introduce unsupported legal conclusions or factual details that were never contained in the underlying record.

The prosecutor must therefore independently establish:

1. the factual basis for each material allegation;
2. the legal basis for each charge;
3. the jurisdictional basis for the proceeding;
4. the accuracy of quoted evidence;
5. the absence of invented facts;
6. and the procedural validity of the document.

The use of GAI may reduce drafting time, but efficiency cannot justify delegating the substantive formulation of criminal allegations to a probabilistic system.

5. Confidentiality and Sensitive Criminal Information

Criminal prosecution involves highly sensitive information, including personal identifiers, medical records, communications, witness information, investigative materials, and evidence concerning victims and suspects.

The use of external GAI systems can create confidentiality concerns when prosecutors transmit sensitive information to third-party platforms.

Formal Opinion 512 specifically identifies confidentiality as an ethical obligation implicated by generative AI.

The relevant issue is broader than traditional attorney-client privilege. Prosecutorial information may be subject to statutory confidentiality obligations, investigative restrictions, privacy law, protective orders, and institutional security requirements.

A prosecutor therefore must determine whether a proposed AI deployment permits the disclosure, retention, processing, or secondary use of sensitive information in a manner compatible with applicable law.

The fact that an AI service is technologically sophisticated does not establish legal authorization to transmit protected criminal-case information to it.

6. Candor Toward the Tribunal

Professional responsibility also imposes duties concerning candor.

Model Rule 3.3 addresses candor toward the tribunal, while Formal Opinion 512 recognizes that GAI use does not diminish the lawyer's obligation to ensure that submissions are accurate and professionally appropriate.

The implication is straightforward: AI-generated content cannot be used to manufacture apparent authority.

If a prosecutor submits an argument containing a citation generated by AI without verification, the responsibility for the inaccurate citation remains with the prosecutor.

The same principle applies to factual summaries. A generated summary cannot be treated as an evidentiary source merely because it accurately resembles the record. The underlying evidence remains the legally relevant material.

7. Traceability and Record Preservation

A further issue concerns traceability.

When AI materially influences a prosecutorial decision, questions may arise concerning:

- which system was used;

- what task it performed;
- what material it processed;
- what output it generated;
- whether the output was modified;
- who reviewed it;
- and what independent verification occurred.

Not every use of AI requires preservation of every prompt or intermediate output. Such a requirement could become disproportionate and operationally burdensome. Nevertheless, where AI materially contributes to a consequential prosecutorial decision, the absence of any record may impair institutional accountability.

The EU AI Act provides a useful comparative reference because its regulatory architecture emphasizes logging, documentation, human oversight, and information necessary for interpreting outputs in high-risk contexts.

This does not mean that every prosecution must automatically adopt the AI Act's documentation regime. Rather, the comparative principle is that consequential AI use should be sufficiently traceable to permit later assessment of responsibility.

8. Disclosure of AI Use

Whether prosecutors must disclose the use of GAI in every case is a more difficult question.

A categorical disclosure rule would be unnecessarily broad. If a prosecutor uses an AI system merely to correct spelling or format an internal administrative document, disclosure would generally have no meaningful relationship to the defendant's procedural rights.

The analysis changes when AI materially influences:

- the identification of evidence;
- the characterization of evidence;
- the generation of factual allegations;
- the selection of legal authorities;

- the preparation of expert-related material;
- or another aspect of the prosecution that bears directly upon the defense's ability to test the case.

The relevant criterion should therefore be **materiality rather than mere technological involvement**.

Where AI use materially affects evidence or reasoning that the prosecution relies upon, information concerning the system's role may become procedurally relevant. The purpose of disclosure would not be to create a generalized right to inspect every technological tool used by the state, but to ensure that the defense can meaningfully challenge the evidentiary or legal basis of the prosecution where AI has materially shaped it.

9. Human Oversight Must Be Substantive

The concept of human oversight should not be reduced to formal approval.

A prosecutor who routinely accepts AI-generated summaries without examining source materials exercises nominal rather than substantive oversight.

Effective oversight requires at least:

- competence to understand the system's relevant limitations;
- authority to reject its output;
- access to the underlying evidence or authority;
- sufficient time to conduct verification;
- and responsibility for the final legal decision.

This approach is consistent with the Council of Europe's broader AI framework. The European Ethical Charter for AI in judicial systems identifies fundamental rights, non-discrimination, quality and security, transparency, fairness, and user control as core principles.

The Council of Europe Framework Convention on AI and Human Rights, Democracy and the Rule of Law further establishes a legally binding international framework

centered on human rights, democracy, and the rule of law, including transparency, oversight, accountability, responsibility, equality, and procedural safeguards.

These standards reinforce a general legal proposition: human involvement is meaningful only when the human decision-maker retains actual authority and capacity to evaluate the technological output.

10. The Limits of Delegated Prosecutorial Judgment

Certain decisions should remain fundamentally attributable to the prosecutor.

These include:

- whether probable cause exists;
- whether charges should be initiated;
- whether evidence is sufficiently reliable to be presented;
- whether disclosure obligations have been satisfied;
- whether a factual allegation is supported;
- whether a legal argument is sustainable;
- whether a plea position is justified;
- and whether a representation to the court is accurate.

AI can assist with these decisions by organizing information or identifying issues. It should not independently determine them.

The distinction is consistent with the broader European regulatory approach. The AI Act's treatment of AI used in judicial contexts recognizes the importance of preserving human decision-making where legal consequences are significant.

The underlying principle is institutionally conservative: legal responsibility follows legal authority.

A prosecutor cannot delegate legal responsibility merely by delegating a cognitive task to a technological system.

Discussion

The analysis demonstrates that the principal legal challenge created by prosecutorial GAI use is not the existence of artificial intelligence itself. The challenge is the transformation of professional workflows in ways that can obscure responsibility.

Traditional criminal procedure assumes identifiable institutional actors. A prosecutor reviews evidence, exercises judgment, makes representations, and bears responsibility for those actions. GAI introduces an additional layer between information and legal judgment.

This layer can improve efficiency but can also produce what may be described as **responsibility diffusion**. When an inaccurate proposition originates from an AI system, the human actor may be tempted to regard the error as technological rather than professional. Criminal procedure cannot accommodate such a transfer of responsibility.

The prosecutor remains the legally competent actor.

The second problem is **verification asymmetry**. Generative systems can produce outputs much faster than human professionals can verify them. This creates an institutional temptation to accept outputs because verification is more expensive than generation. In criminal proceedings, however, the relevant standard cannot be determined by comparative convenience.

The state possesses coercive authority. Efficiency therefore cannot override accuracy and procedural fairness.

The third problem is **automation-induced confidence**. A professionally formatted memorandum or factual summary may appear more authoritative than its underlying foundation warrants. The legal system must therefore distinguish presentation quality from evidentiary reliability.

The fourth problem concerns disclosure. AI-related information should become legally relevant when it materially affects the ability of the defense to challenge the prosecution’s case. This approach avoids both extremes: treating every use of AI as discoverable regardless of relevance, or treating AI use as categorically immune from scrutiny.

The comparative materials support a common direction. The Council of Europe emphasizes transparency, fairness, accountability, and human control. The EU AI Act formalizes human oversight and documentation requirements for high-risk contexts. The ABA applies traditional professional duties to GAI use rather than creating a separate ethical universe for artificial intelligence.

The resulting legal position is therefore evolutionary rather than revolutionary.

Existing criminal procedure already contains the essential principles. What is required is their disciplined application to a new technological environment.

Limitations

This analysis is primarily doctrinal and comparative. It does not provide empirical measurements of prosecutorial AI adoption, error rates, hallucination frequencies, or jurisdiction-specific implementation practices.

The legal treatment of GAI also remains dynamic. National legislation, professional rules, prosecutorial policies, and judicial decisions continue to develop. The AI Act itself contains differentiated obligations whose application depends on the classification and context of a particular AI system.

The comparative analysis should therefore not be understood as asserting that one jurisdiction’s rules automatically apply in another. Instead, the selected jurisdictions provide normative and doctrinal reference points for identifying principles capable of informing broader criminal-procedure analysis.

A further limitation concerns the distinction between generative AI and other AI systems. Predictive analytics, biometric identification, document classification, large language models, and automated discovery tools present different technical and legal characteristics. This article focuses specifically on generative systems used in prosecutorial legal and evidentiary workflows.

Conclusions

Generative artificial intelligence can legitimately assist prosecutorial work, but its use must remain subordinate to professional responsibility, evidentiary reliability, and human legal judgment.

The first conclusion is that **AI output is not equivalent to legal authority or evidence**. A generated proposition acquires legal significance only when a competent human verifies its source and accuracy.

The second conclusion is that **professional responsibility cannot be delegated to technology**. Prosecutors remain responsible for charging decisions, factual allegations, legal representations, disclosure obligations, and submissions to tribunals regardless of whether AI contributed to their preparation.

The third conclusion is that **human oversight must be substantive**. Effective oversight requires competence, authority, access to underlying materials, and the practical capacity to reject an AI-generated conclusion.

The fourth conclusion is that **disclosure should be governed by materiality**. The mere use of AI for administrative assistance should not automatically generate a procedural disclosure obligation. However, when AI materially contributes to the identification, interpretation, characterization, or presentation of evidence or legal reasoning on which the prosecution relies, the system's role may become relevant to the defense's ability to test the case.

The fifth conclusion is that **confidentiality must remain technologically neutral**. Prosecutorial obligations concerning sensitive information do not disappear because information is transmitted to an AI system rather than to a conventional service provider.

The sixth conclusion is that **traceability becomes increasingly important as AI assumes greater functional significance**. Where AI materially influences consequential prosecutorial decisions, adequate records should exist to permit institutional review of how the output was generated, evaluated, and ultimately used, subject to applicable confidentiality and data-protection requirements.

The seventh conclusion is that criminal justice should not treat generative AI as an autonomous legal decision-maker. The more consequential the decision, the stronger the justification for maintaining identifiable human responsibility.

The resulting doctrinal position is therefore neither technological prohibition nor technological deference. It is controlled professional integration.

Generative AI may accelerate legal work, organize complex information, assist discovery, facilitate research, and support drafting. It may not, however, become the hidden decision-maker behind a criminal prosecution.

The governing principle is simple but legally consequential: **the state may use artificial intelligence to assist prosecutorial judgment, but it may not use artificial intelligence to obscure who exercised that judgment or upon what verified foundation it was exercised.**

References

American Bar Association. (2024). *Formal Opinion 512: Generative artificial intelligence tools*. Standing Committee on Ethics and Professional Responsibility.

American Bar Association. (2024). *Model Rules of Professional Conduct: Rule 3.8—Special responsibilities of a prosecutor*.

American Bar Association. (2024). *Model Rules of Professional Conduct: Rule 3.3—Candor toward the tribunal*.

Council of Europe, European Commission for the Efficiency of Justice. (2018). *European Ethical Charter on the use of artificial intelligence in judicial systems and their environment*.

Council of Europe. (2024). *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law* (CETS No. 225).

European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union.

Federal Rules of Criminal Procedure. (2026). *Rule 16: Discovery and inspection*. United States Courts.

United States Department of Justice. (2024). *Artificial Intelligence and Criminal Justice: Final Report*. Office of Legal Policy.